An Outer Bound to the Capacity Region of the Broadcast Channel

Chandra Nair, *Member, IEEE,* and Abbas El Gamal, *Fellow, IEEE.*

arXiv:cs/0605105v2 [cs.IT] 5 Oct 2006

***Abstract*— An outer bound to the capacity region of the two-receiver discrete memoryless broadcast channel is given. The outer bound is tight for all cases where the capacity region is known. When specialized to the case of no common information, this outer bound is contained in the Körner-Marton outer bound. This containment is shown to be strict for the *binary skew-symmetric* broadcast channel. Thus, this outer bound is in general tighter than all other known outer bounds.**

***Index Terms*— broadcast channel, capacity, outer bound**

1. Introduction

We consider a discrete memoryless (DM) broadcast channel where the sender wishes to communicate common as well as separate messages to two receivers [3]. Formally, the channel consists of an input alphabet $\mathcal{X}$, output alphabets $\mathcal{Y}$ and $\mathcal{Z}$, and a probability transition function $p(y,z|x)$. A $((2^{nR_0}, 2^{nR_1}, 2^{nR_2}), n)$ code for this channel consists of (i) three messages (M_0, M_1, M_2) uniformly distributed over $[1, 2^{nR_0}] \times [1, 2^{nR_1}] \times [1, 2^{nR_2}]$, (ii) an encoder that assigns a codeword $x^n(m_0, m_1, m_2)$, for each message triplet $(m_0, m_1, m_2) \in [1, 2^{nR_0}] \times [1, 2^{nR_1}] \times [1, 2^{nR_2}]$, and (iii) two decoders, one that maps each received y^n sequence into an estimate $(\hat{m}_0, \hat{m}_1) \in [1, 2^{nR_0}] \times [1, 2^{nR_1}]$ and another that maps each received z^n sequence into an estimate $(\hat{\hat{m}}_0, \hat{m}_2) \in [1, 2^{nR_0}] \times [1, 2^{nR_2}]$.

The probability of error is defined as

$$\begin{aligned} P_e^{(n)} = \mathrm{P}(\hat{M}_0 \neq M_0 \text{ or } \hat{\hat{M}}_0 \neq M_0 \text{ or } \hat{M}_1 \neq M_1 \\ \text{or } \hat{M}_2 \neq M_2). \end{aligned}$$

A rate tuple (R_0, R_1, R_2) is said to be achievable if there exists a sequence of $((2^{nR_0}, 2^{nR_1}, 2^{nR_2}), n)$ codes with $P_e^{(n)} \to 0$. The capacity region of the broadcast channel is the closure of the set of achievable rates.

The capacity region for this channel is known only for some classes, including the degraded [2], [7], [1], less noisy[10], more capable[6], deterministic[11], [13] and semi-deterministic channels[8]. Additionally, general inner bounds by Cover [4], van der Meulen [15] and Marton [12] and outer bounds by Körner and Marton [12] and Sato [14] have been established. Furthermore, the Körner and Marton [12] outer bound was found to be tight for all cases where capacity is known.

In this paper we introduce an outer bound on the capacity region of the DM broadcast channel based on results in [6] and show that it is strictly tighter than existing outer bounds. The outer bound is presented in the next section. In Section 3, the outer bound is specialized to the case of no common information. In Section 3-C, it is shown that that when there is no common information, our outer bound is contained in the Körner-Marton bound and in Section 4 it is shown that this containment is strict.

2. Outer bound

The following is an outer bound to the capacity region of the two-receiver DM broadcast channel.

Theorem 2.1: The set of rate triples (R_0, R_1, R_2) satisfying

$$\begin{aligned} R_0 &\leq \min\{I(W;Y), I(W;Z)\}, \\ R_0 + R_1 &\leq I(U,W;Y), \\ R_0 + R_2 &\leq I(V,W;Z), \\ R_0 + R_1 + R_2 &\leq I(U,W;Y) + I(V;Z|U,W), \\ R_0 + R_1 + R_2 &\leq I(V,W;Z) + I(U;Y|V,W), \end{aligned}$$

for some joint distribution of the form $p(u,v,w,x) = p(u)p(v)p(w|u,v)p(x|u,v,w)$ constitutes an outer bound to the capacity region for the DM broadcast channel.

Proof: The arguments are essentially the same as those used in the converse proof for the more capable

Chandra Nair is a post-doctoral researcher with the Theory Group at Microsoft Research.

Abbas El Gamal is a professor with the Electrical Engineering Dept. at Stanford University.

broadcast channel class [6]. Observe that

$$
\begin{aligned}
nR_0 &= H(M_0) \\
&= H(M_0|Y^n) + I(M_0;Y^n) \\
&\overset{(a)}{\le} n\lambda_{0n} + \sum_{i=1}^{n} \left(H(Y_i|Y^{i-1}) - H(Y_i|M_0,Y^{i-1})\right) \\
&\overset{(b)}{\le} n\lambda_{0n} + \sum_{i=1}^{n} \left(H(Y_i) - H(Y_i|M_0,Y^{i-1},Z_{i+1}^n)\right),
\end{aligned}
$$

where (a) follows by Fano's inequality and (b) follows from the fact that conditioning decreases entropy. Now defining the random variable $W_i = (M_0, Y^{i-1}, Z_{i+1}^n)$, we obtain

$$
\begin{aligned}
nR_0 &\le n\lambda_{0n} + \sum_{i=1}^{n} (H(Y_i) - H(Y_i|W_i)) \\
&= n\lambda_{0n} + \sum_{i=1}^{n} I(Y_i;W_i).
\end{aligned} \tag{2.1}
$$

In a similar fashion observe that

$$
\begin{aligned}
nR_0 &= H(M_0) \\
&= H(M_0|Z^n) + I(M_0;Z^n) \\
&\le n\lambda_{1n} + \sum_{i=1}^{n} \left(H(Z_i|Z_{i+1}^n) - H(Z_i|M_0,Z_{i+1}^n)\right) \\
&\le n\lambda_{1n} + \sum_{i=1}^{n} \left(H(Z_i) - H(Z_i|M_0,Y^{i-1},Z_{i+1}^n)\right) \\
&= n\lambda_{1n} + \sum_{i=1}^{n} I(Z_i;W_i).
\end{aligned} \tag{2.2}
$$

Now, consider

$$
\begin{aligned}
&n(R_0+R_1) \\
&= H(M_0,M_1) \\
&= H(M_0,M_1|Y^n) + I(M_0,M_1;Y^n) \\
&\le n\lambda_{2n} + \sum_{i=1}^{n} \left(H(Y_i|Y^{i-1}) - H(Y_i|M_0,M_1,Y^{i-1})\right) \\
&\le n\lambda_{2n} + \sum_{i=1}^{n} \left(H(Y_i) - H(Y_i|M_0,M_1,Y^{i-1},Z_{i+1}^n)\right) \\
&\le n\lambda_{2n} + \sum_{i=1}^{n} I(Y_i;U_i,W_i),
\end{aligned} \tag{2.3}
$$

where we define the random variable $U_i = M_1$ for all i.

In a similar fashion

$$
n(R_0+R_2) \le n\lambda_{3n} + \sum_{i=1}^{n} I(Z_i;V_i,W_i), \tag{2.4}
$$

where $V_i = M_2$ for all i.

Lastly, consider

$$
\begin{aligned}
&n(R_0+R_1+R_2) \\
&= H(M_0,M_1,M_2) \\
&= H(M_0,M_1) + H(M_2|M_0,M_1) \\
&\le n\lambda_{4n} + I(M_0,M_1;Y^n) + I(M_2;Z^n|M_0,M_1)
\end{aligned}
$$

$$
\begin{aligned}
&= n\lambda_{4n} + \sum_{i=1}^{n} I(M_0,M_1;Y_i|Y^{i-1}) \\
&\quad + \sum_{i=1}^{n} I(M_2;Z_i|M_0,M_1,Z_{i+1}^n).
\end{aligned} \tag{2.5}
$$

Note that

$$
\begin{aligned}
&\sum_{i=1}^{n} I(M_0,M_1;Y_i|Y^{i-1}) \\
&\le \sum_{i=1}^{n} I(M_0,M_1,Y^{i-1};Y_i) \\
&= \sum_{i=1}^{n} I(M_0,M_1,Y^{i-1},Z_{i+1}^n;Y_i) \\
&\quad - \sum_{i=1}^{n} I(Z_{i+1}^n;Y_i|M_0,M_1,Y^{i-1}).
\end{aligned} \tag{2.6}
$$

And further,

$$
\begin{aligned}
&\sum_{i=1}^{n} I(M_2;Z_i|M_0,M_1,Z_{i+1}^n) \\
&\le \sum_{i=1}^{n} I(M_2,Y^{i-1};Z_i|M_0,M_1,Z_{i+1}^n) \\
&= \sum_{i=1}^{n} I(Y^{i-1};Z_i|M_0,M_1,Z_{i+1}^n) \\
&\quad + \sum_{i=1}^{n} I(M_2;Z_i|M_0,M_1,Z_{i+1}^n,Y^{i-1}).
\end{aligned} \tag{2.7}
$$

Combining equations (2.5), (2.6), (2.7), we obtain

$$
\begin{aligned}
&n(R_0+R_1+R_2) \\
&\le n\lambda_{4n} + \sum_{i=1}^{n} I(M_0,M_1,Y^{i-1},Z_{i+1}^n;Y_i) \\
&\quad - \sum_{i=1}^{n} I(Z_{i+1}^n;Y_i|M_0,M_1,Y^{i-1}) \\
&\quad + \sum_{i=1}^{n} I(Y^{i-1};Z_i|M_0,M_1,Z_{i+1}^n) \\
&\quad + \sum_{i=1}^{n} I(M_2;Z_i|M_0,M_1,Z_{i+1}^n,Y^{i-1})
\end{aligned}
$$

$$\stackrel{(d)}{=} n\lambda_{4n} + \sum_{i=1}^{n} I(M_0, M_1, Y^{i-1}, Z_{i+1}^n; Y_i)$$
$$+ \sum_{i=1}^{n} I(M_2; Z_i | M_0, M_1, Z_{i+1}^n, Y^{i-1})$$
$$= n\lambda_{4n} + \sum_{i=1}^{n} \left(I(U_i, W_i; Y_i) + I(V_i; Z_i | U_i, W_i) \right). \tag{2.8}$$

The equality (d) follows from the well known Csiszar identity (see Lemma 7 in [5]):

$$\sum_{i=1}^{n} I(\mathbf{Y}^{i-1}; Z_i | \mathbf{Z}_{i+1}^n) = \sum_{i=1}^{n} I(\mathbf{Z}_{i+1}^n; Y_i | \mathbf{Y}^{i-1}).$$

In a very similar fashion, we can also obtain

$$n(R_0 + R_1 + R_2) \leq n\lambda_{5n} + \sum_{i=1}^{n} \left(I(V_i, W_i; Z_i) + I(U_i; Y_i | V_i, W_i) \right). \tag{2.9}$$

Define the time sharing random variable Q to be independent of $M_0, M_1, M_2, X^n, Y^n, Z^n$, and uniformly distributed over $\{1, 2, .., n\}$ and define $W = (Q, W_Q)$, $U = U_Q$, $V = V_Q$, $X = X_Q$, $Y = Y_Q$, $Z = Z_Q$.

Clearly we have

$$\begin{aligned} nR_0 &\leq n\lambda_{0n} + \sum_{i=1}^{n} I(W_i; Y_i) \\ &= n\lambda_{0n} + nI(W; Y|Q) \\ &\leq n\lambda_{0n} + nI(W; Y). \end{aligned} \tag{2.10}$$

Similarly,

$$\begin{aligned} nR_0 &\leq n\lambda_{1n} + nI(W; Z) \\ n(R_0 + R_1) &\leq n\lambda_{2n} + nI(U, W; Y) \\ n(R_0 + R_2) &\leq n\lambda_{3n} + nI(V, W; Z). \end{aligned} \tag{2.11}$$

Further,

$$\begin{aligned} &n(R_0 + R_1 + R_2) \\ &\leq n\lambda_{4n} + \sum_{i=1}^{n} \left(I(U_i, W_i; Y_i) + I(V_i; Z_i | U_i, W_i) \right) \\ &= n\lambda_{4n} + nI(U, W; Y|Q) + nI(V; Z|U, W, Q) \\ &= n\lambda_{4n} + nI(U, W; Y|Q) + nI(V; Z|U, W) \\ &\leq n\lambda_{4n} + nI(U, W; Y) + nI(V; Z|U, W), \end{aligned} \tag{2.12}$$

and similarly

$$\begin{aligned} &n(R_0 + R_1 + R_2) \\ &\quad \leq n\lambda_{5n} + nI(V, W; Z) + nI(U; Y|V, W). \end{aligned}$$

The independence of the messages M_1 and M_2 implies the independence of the auxiliary random variables U and V as specified.

Since the probability of error is assumed to tend to zero, $\lambda_{0n}, \lambda_{1n}, \lambda_{2n}, \lambda_{3n}, \lambda_{4n}$, and λ_{5n} also tend to zero as $n \to \infty$. This completes the proof. ■

3. Outer Bound with No Common Information

Note that the outer bound given in Theorem 2.1 immediately leads to the following outer bound for the case when there is no common information, i.e., $R_0 = 0$.

The set of all rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(U, W; Y), \\ R_2 &\leq I(V, W; Z), \\ R_1 + R_2 &\leq I(U, W; Y) + I(V; Z|U, W) \\ R_1 + R_2 &\leq I(V, W; Z) + I(U; Y|V, W), \end{aligned} \tag{3.1}$$

for some joint distribution of the form $p(u, v, w, x) = p(u)p(v)p(w|u, v)p(x|u, v, w)$ constitutes an outer bound on the capacity of the DM broadcast channel with no common information.

The following theorem gives a *possibly* weaker outer bound that we consider for the rest of the paper.

Theorem 3.1: Consider the DM broadcast channel with no common information. The set of rate pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq I(U; Y), \\ R_2 &\leq I(V; Z), \\ R_1 + R_2 &\leq I(U; Y) + I(V; Z|U), \\ R_1 + R_2 &\leq I(V; Z) + I(U; Y|V), \end{aligned}$$

for some choice of joint distributions $p(u, v, x) = p(u, v)p(x|u, v)$ constitutes an outer bound to the capacity region for the DM broadcast channel with no common information.

Proof: This follows by redefining U as (U, W) and V as (V, W) in equation (3.1). ■

In the following subsections we prove results that aid in the evaluation of the above outer bound.

A. X Deterministic Function of U, V suffices

Denote by $\mathcal{C}$ the outer bound in Theorem 3.1 and let $\mathcal{C}_d$ be the same bound but with X restricted to be a deterministic function of U, V, i.e. $\mathrm{P}(X = x|U = u, V = v) \in \{0, 1\}$ for all (u, v, x). We now show that these two bounds are identical.

Lemma 3.2: $\mathcal{C} = \mathcal{C}_d$.

Before we prove this lemma, note that it suffices to show that $\mathcal{C} \subset \mathcal{C}_d$. Our method of proof is as follows: For

every $p(u,v), p(x|u,v)$ we will construct random variables U^*, V^*, X^* where X^* is a deterministic function of U^*, V^* such that the region described by U^*, V^*, X^* will contain the region described by U, V, W.

Let $\mathrm{P}(U = u, V = v) = p_{uv}$ and $\mathrm{P}(X = x|U = u, V = v) = \delta_{uv}^x \in \{0, 1\}$. Without loss of generality, assume that $\mathcal{X} = \{0, 1, .., m-1\}$. Now, construct random variables U^*, V^* having cardinalities $m\|U\|, m\|V\|$ as follows: Split each value u taken by U into m values $u_0, \ldots, u_{m-1}$ and each value v taken by V into m values $v_0, \ldots, v_{m-1}$. Let

$$\begin{aligned}
&\mathrm{P}(U^* = u_i, V^* = v_j) \\
&\qquad = \frac{1}{m}\mathrm{P}(U = u, V = v, X = (i-j)_m), \\
&\mathrm{P}(X^* = k|U^* = u_i, V^* = v_j) \\
&\qquad = \begin{cases} 1 & \text{if } k = (i-j)_m \\ 0 & \text{otherwise,} \end{cases}
\end{aligned} \tag{3.2}$$

where $(l)_m$ is the remainder of l/m (the $\mod$ operation).

We will need the following facts.

Lemma 3.3: The following hold:

(i) $\mathrm{P}(U^* = u_i) = \frac{1}{m}\mathrm{P}(U = u)$ for $0 \le i \le m-1$.
(ii) $\mathrm{P}(V^* = v_i) = \frac{1}{m}\mathrm{P}(V = v)$ for $0 \le i \le m-1$.
(iii) $\mathrm{P}(X^* = k|U^* = u_i) = \mathrm{P}(X = k|U = u)$ for $0 \le k, i \le m-1$.
(iv) $\mathrm{P}(X^* = k|V^* = v_i) = \mathrm{P}(X = k|V = v)$ for $0 \le k, i \le m-1$.

Proof: Observe that

$$\begin{aligned}
&\mathrm{P}(U^* = u_i) \\
&\quad = \sum_{v \in \mathcal{V}} \sum_{j=1}^{m} \mathrm{P}(U^* = u_i, V^* = v_j) \\
&\quad = \sum_{v \in \mathcal{V}} \sum_{j=1}^{m} \frac{1}{m}\mathrm{P}(U = u, V = v, X = (i-j)_m) \\
&\quad = \sum_{v \in \mathcal{V}} \frac{1}{m}\mathrm{P}(U = u, V = v) \\
&\quad = \frac{1}{m}\mathrm{P}(U = u).
\end{aligned}$$

Proof of (ii) follows similarly. To show (iii), consider

$$\begin{aligned}
&\mathrm{P}(X^* = k|U^* = u_i) \\
&\quad = \sum_{v \in \mathcal{V}} \sum_{j=1}^{m} \mathrm{P}(X^* = k, V^* = v_j|U^* = u_i) \\
&\quad \overset{(a)}{=} \sum_{v \in \mathcal{V}} \mathrm{P}(X^* = k, V^* = v_{(i-k)_m}|U^* = u_i) \\
&\quad = \sum_{v \in \mathcal{V}} \mathrm{P}(V^* = v_{(i-k)_m}|U^* = u_i) \\
&\quad = \sum_{v \in \mathcal{V}} \frac{1}{m} \frac{\mathrm{P}(X = k, V = v, U = u)}{P(U^* = u_i)} \\
&\quad \overset{(b)}{=} \sum_{v \in \mathcal{V}} \mathrm{P}(X = k, V = v|U = u) \\
&\quad = \mathrm{P}(X = k|U = u),
\end{aligned}$$

where (a) follows from the fact that the rest of the terms are zero by construction and (b) follows from (i) using the fact that $\mathrm{P}(U^* = u_i) = \frac{1}{m}\mathrm{P}(U = u)$. The proof of (iv) follows similarly. ■

The following corollary follows from the above lemma, the fact that X^* is a deterministic function of (U^*, V^*), and the fact that $(U^*, V^*) \to X^* \to (Y^*, Z^*), (U, V) \to X \to (Y, Z)$ form Markov chains with $p(y^*, z^*|x^*) = p(y, z|x)$. The proofs are straightforward and are therefore omitted.

Corollary 3.4: The following hold:

(i) $P(X^* = i) = P(X = i)$ for $0 \le i \le m-1$.
(ii) $H(Y^*|U^*) = H(Y|U)$.
(iii) $H(Z^*|U^*) = H(Z|U)$.
(iv) $H(Y^*|V^*) = H(Y|V)$.
(v) $H(Z^*|V^*) = H(Z|V)$.
(vi) $H(Y^*|U^*, V^*) = H(Y^*|X^*)$
$= H(Y|X) \le H(Y|U, V)$.
(vii) $H(Z^*|U^*, V^*) = H(Z^*|X^*)$
$= H(Z|X) \le H(Z|U, V)$.

We are now ready to prove Lemma 3.2

Proof of Lemma 3.2: Corollary 3.4 implies that

$$\begin{aligned}
I(U; Y) &= I(U^*; Y^*), \\
I(V; Z) &= I(V^*; Z^*), \\
I(U; Y|V) &\le I(U^*; Y^*|V^*), \\
I(V; Z|U) &\le I(V^*; Z^*|U^*), \\
I(X; Y|V) &= I(X^*; Y^*|V^*), \\
I(X; Z|U) &= I(X^*; Z^*|U^*).
\end{aligned} \tag{3.3}$$

Thus $\mathcal{C} \subset \mathcal{C}_d$, which completes the proof of Lemma 3.2.

Thus the outer bound in Theorem 3.1 can be re-expressed as follows.

Lemma 3.5: The set of rate pairs satisfying

$$\begin{aligned} R_1 &\le I(U;Y), \\ R_2 &\le I(V;Z), \\ R_1+R_2 &\le I(U;Y)+I(X;Z|U), \\ R_1+R_2 &\le I(V;Z)+I(X;Y|V), \end{aligned}$$

for some distribution $p(u,v,x) = p(u,v)p(x|u,v)$, where $p(x|u,v) \in \{0,1\}$, constitute an outer bound on the DM broadcast channel with no common information.

Remark 3.6: Note that the constraint $p(x|u,v) \in \{0,1\}$ while useful for evaluating the region, can be removed from the definition, since as before, for any (U,V,X) one can construct random variables (U^*,V^*,X^*) according to equation (3.2) and by equation (3.3), the region (R_1,R_2) evaluated using (U,V,X) is identical to that evaluated using (U^*,V^*,X^*).

B. *Cardinality bounds on U and V*

We now establish bounds on the cardinality of U and V. From Remark 3.6, we know that $p(x|u,v)$ can be arbitrary.

Fact 3.7: Given $p(u), p(x|u), p(v), p(x|v)$, if $p(x)$ is consistent, i.e.,

$$\sum_{u\in\mathcal{U}} p(X=x|u)p(u) = \sum_{v\in\mathcal{V}} p(X=x|v)p(v)$$

for every $x \in \mathcal{X}$, then there exist $p(u,v)$ and $p(x|u,v)$ that are consistent with $p(u), p(x|u), p(v), p(x|v)$.

Remark 3.8: A canonical way to generate such a joint triple is to generate X according to $p(x)$ and then generate U, V conditionally independent of X according to $p(u|x)$ and $p(v|x)$.

Now for any $U \to X \to (Y,Z)$, using standard arguments from [1], there exists a (U^*,X^*) with $\|U^*\| \le \|X\|+2$, such that $I(U;Y) = I(U^*;Y^*)$ and $I(X;Z|U) = I(X^*;Z^*|U^*)$. Similarly, there exists a V^* with $\|V^*\| \le \|X\|+2$, such that $I(V;Z) = I(V^*;Z^*)$ and $I(X;Y|V) = I(X^*;Y^*|V^*)$. From Fact 3.7, it follows that there exists a triple (U^*,V^*,X^*) consistent with the pairs (U^*,X^*) and (V^*,X^*). Thus we can assume that $\|U\| \le \|X\|+2, \ \ \|V\| \le \|X\|+2$.

C. *Comparison to Körner-Marton outer bound*

The outer bound of Körner and Marton [12] is given by $\mathcal{O} = \mathcal{O}_y \cap \mathcal{O}_z$, where $\mathcal{O}_y$ is the set of rate pairs (R_1,R_2) satisfying

$$\begin{aligned} R_1 &\le I(X;Y), \\ R_2 &\le I(V;Z), \\ R_1+R_2 &\le I(V;Z)+I(X;Y|V), \end{aligned}$$

for some distribution $p(v)p(x|v)$, and $\mathcal{O}_z$ is the set of rate pairs (R_1,R_2) satisfying

$$\begin{aligned} R_2 &\le I(X;Z), \\ R_1 &\le I(U;Y), \\ R_1+R_2 &\le I(U;Y)+I(X;Z|V), \end{aligned}$$

for some distribution $p(u)p(x|u)$.

From Lemma 3.5, it is clear that $\mathcal{C} \subset \mathcal{O}_y$ and $\mathcal{C} \subset \mathcal{O}_z$. Hence

$$\mathcal{C} \subset \mathcal{O} = \mathcal{O}_y \cap \mathcal{O}_z$$

and $\mathcal{C}$ is in general contained in the Körner-Marton outer bound. In the following section, we show that the containment is strict for the *binary skew-symmetric* broadcast channel.

4. Binary Skew-Symmetric Channel

Consider the Binary Skew-Symmetric Channel (BSSC) shown in Figure 1, which was studied by Hajek and Pursley [9]. For the rest of the paper we assume that $p = \frac{1}{2}$, though a similar analysis can be carried out for any other choice of p.

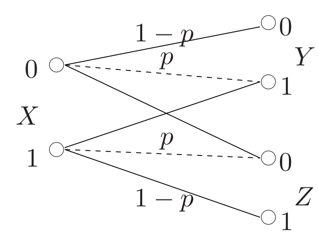

Fig. 1. Binary Skew Symmetric Channel

In [9], the Cover-van der Meulen achievable rate region for the DM broadcast channel, $\mathcal{D}$, was evaluated for the Binary Skew-Symmetric broadcast channel (BSSC) with private messages only. The resulting coding scheme has the following intuitive interpretation, which we denote by *randomized time-sharing*. Observe that if $X = 0$ is sent, it is received error free by Z, but completely noisy by Y. Conversely, if $X = 1$ is sent, it is received error free by Y, but completely noisy by Z. This suggests that a time-sharing scheme, where transmission time is divided between the two user before communication commences, is optimal. It turned out that higher rates can be achieved by performing randomized time-sharing, instead. This is done via a *common information* random variable W, which specifies the locations of the symbols in the received sequence corresponding to each user's *private message*. Each receiver first decodes W to find out which part of the received sequence corresponds

to its private message, then proceeds to decode its private message. Using standard random coding and joint typicality decoding arguments, it is can be shown that any (R_1,R_2) satisfying the conditions

$$\begin{aligned}
R_1 &\leq \min\{I(W;Y), I(W;Z)\}\\
&\quad + \mathrm{P}(W=0)I(X;Y|W=0),\\
R_2 &\leq \min\{I(W;Y), I(W;Z)\}\\
&\quad + \mathrm{P}(W=1)I(X;Z|W=1),\\
R_1+R_2 &\leq \min\{I(W;Y), I(W;Z)\}\\
&\quad + \mathrm{P}(W=0)I(X;Y|W=0)\\
&\quad + \mathrm{P}(W=1)I(X;Z|W=1),
\end{aligned}$$

for some $p(w)p(x|w)$, is achievable.

The line segment joining $(R_1,R_2) = (0.2411..,0.1204..)$ to $(R_1,R_2) = (0.1204..,0.2411..)$ is achieved by the following choice of (W,X). Let $\alpha = 0.5-\sqrt{105}/30 \approx 0.1584$, then

$$\begin{aligned}
\mathrm{P}(W=0) = \mathrm{P}(W=1) &= 0.5,\\
\mathrm{P}(X=0|W=0) &= \alpha,\\
\mathrm{P}(X=1|W=1) &= \alpha.
\end{aligned}$$

It is not difficult to see that the line segment joining $(R_1,R_2) = (0.2411..,0.1204..)$ and $(R_1,R_2) = (0.1204..,0.2411..)$ lies on the boundary of this region. Note that on this line segment $R_1+R_2 = 0.3616....$

We now show that the region $\mathcal{C}$, described by the outer bound, is strictly larger than the Cover-van der Meulen region $\mathcal{D}$ and strictly smaller than the Körner-Marton outer bound.

Claim 4.1: The line segment connecting $(R_1,R_2) = (0.2280..,0.1431..)$ to $(R_1,R_2) = (0.1431..,0.2280..)$ lies on the boundary of $\mathcal{C}$.

Proof: Note that from Lemma 3.5, the sum rate is bounded by

$$\begin{aligned}
R_1+R_2 &\leq \frac{1}{2}(I(U;Y)+I(X;Z|U))\\
&\quad + \frac{1}{2}(I(V;Z)+I(X;Y|V)).
\end{aligned}$$

We proceed to maximize the RHS of the above inequality over $p(u,v,x)$. Assume that (U_o,V_o,X_o) maximizes the sum rate and let

$$\begin{aligned}
R_m &= \frac{1}{2}(I(U_o;Y_o)+I(X_o;Z_o|U_o))\\
&\quad + \frac{1}{2}(I(V_o;Z_o)+I(X_o;Y_o|V_o)).
\end{aligned}$$

Consider a triple (U',V',X) with $\mathcal{U}'=\mathcal{V}, \mathcal{V}'=\mathcal{U}$, such that

$$\begin{aligned}
&\mathrm{P}(U'=u',V'=v')\\
&\quad = \mathrm{P}(U_o=v',V_o=u'),\\
&\mathrm{P}(X=x|U'=u',V'=v')\\
&\quad = \mathrm{P}(X=1-x|U_o=v',V_o=u').
\end{aligned} \tag{4.1}$$

By the symmetry of the channel,

$$\begin{aligned}
I(U';Y') &= I(V_o;Z_o),\\
I(V';Z') &= I(U_o;Y_o),\\
I(X';Z'|U') &= I(X_o;Y_o|V_o),\\
I(X';Y'|V') &= I(X_o;Z_o|U_o).
\end{aligned}$$

Therefore,

$$\begin{aligned}
R_m &= \frac{1}{2}(I(U';Y')+I(X';Z'|U'))\\
&\quad + \frac{1}{2}(I(V';Z')+I(X';Y'|V')).
\end{aligned}$$

Let $Q \in \{1,2\}$ be an independent random variable that takes values 1 or 2 with equal probability and define $U^* = (\tilde{U},Q)$ and $V^* = (\tilde{V},Q)$ as $(Q=1,\tilde{U},\tilde{V},X) \sim (U_o,V_o,X)$ and $(Q=2,\tilde{U},\tilde{V},X) \sim (U',V',X)$, respectively. Then

$$\begin{aligned}
&\mathrm{P}(X=x|\tilde{U}=u,\tilde{V}=v,Q=1)\\
&\quad = \mathrm{P}(X=x|U_o=u,V_o=v),\\
&\mathrm{P}(X=x|\tilde{U}=u,\tilde{V}=v,Q=2)\\
&\quad = \mathrm{P}(X'=x|U'=u,V'=v).
\end{aligned}$$

Observe that

$$\begin{aligned}
&I(X^*;Y^*|V^*)\\
&\quad = \frac{1}{2}(I(X_o;Y_o|V_o)+I(X';Y'|V')),
\end{aligned}$$

$$\begin{aligned}
&I(U^*;Y^*)\\
&\quad = H(Y^*)-H(Y^*|U^*)\\
&\quad = H(Y^*)-\frac{1}{2}(H(Y_o|U_o)+H(Y'|U')\\
&\quad \overset{(a)}{\geq} \frac{1}{2}(H(Y_o)+H(Y'))-\frac{1}{2}(H(Y_o|U_o)+H(Y'|U'))\\
&\quad = \frac{1}{2}(I(U_o;Y_o)+I(U';Y')),
\end{aligned}$$

where (a) follows by the concavity of the entropy function.

Similarly

$$\begin{aligned}
I(X^*;Z^*|U^*) &= \frac{1}{2}(I(X_o;Z_o|U_o)+I(X';Z'|U')),\\
I(V^*;Z^*) &\geq \frac{1}{2}(I(V_o;Z_o)+I(V';Z')).
\end{aligned}$$

Therefore,

$$R_m \leq I(U^*;Y)+I(X;Z|U^*)+I(V^*;Z)+I(X;Y|V^*).$$

Now, by the construction of (U^*, V^*, X^*), $\mathrm{P}(X^* = 1) = 0.5$. Thus to compute R_m, it suffices to consider X such that $\mathrm{P}(X = 1) = 0.5$.

Using standard optimization techniques, it is not difficult to see that the following (U, X) and (V, X) maximize the terms $I(U;Y)+I(X;Z|U)$ and $I(V;Z)+I(X;Y|V)$, respectively, subject to $\mathrm{P}(X = 1) = 0.5$. As before, let $\alpha = 0.5 - \sqrt{105}/30 \approx 0.1584$. Then a set of maximizing pairs $\mathrm{P}(U,X)$ and $\mathrm{P}(V,X)$ can be described by

$$\begin{aligned}
\mathrm{P}(U=0) &= \frac{0.5}{1-\alpha}, & \mathrm{P}(U=1) &= \frac{0.5-\alpha}{1-\alpha},\\
\mathrm{P}(X=1|U=0) &= \alpha, & \mathrm{P}(X=1|U=1) &= 1,\\
\mathrm{P}(V=0) &= \frac{0.5}{1-\alpha}, & \mathrm{P}(V=1) &= \frac{0.5-\alpha}{1-\alpha},\\
\mathrm{P}(X=0|V=0) &= \alpha, & \mathrm{P}(X=0|V=1) &= 1.
\end{aligned}$$

Substituting these values, we obtain

$$\begin{aligned}
R_1 + R_2 &\leq \frac{1}{2}(I(U;Y) + I(X;Z|U))\\
&\quad + \leq \frac{1}{2}(I(V;Z) + I(X;Y|V))\\
&\leq 0.3711...
\end{aligned}$$

We now show that this bound on the sum rate is tight.

As before, let $\alpha = 0.5-\sqrt{105}/30 \approx 0.1584$. Consider the following (U, V, X)

$$\begin{aligned}
\mathrm{P}(U=0,V=0) &= \frac{\alpha}{1-\alpha},\\
\mathrm{P}(X=1|U=0,V=0) &= 0.5,\\
\mathrm{P}(U=0,V=1) &= \frac{0.5-\alpha}{1-\alpha},\\
\mathrm{P}(X=1|U=0,V=1) &= 0,\\
\mathrm{P}(U=1,V=0) &= \frac{0.5-\alpha}{1-\alpha},\\
\mathrm{P}(X=1|U=1,V=0) &= 1.
\end{aligned}$$

The region evaluated by this (U, V, X) is given by all rate pairs (R_1, R_2) satisfying

$$\begin{aligned}
R_1 &\leq I(U;Y) = 0.2280..,\\
R_2 &\leq I(V;Z) = 0.2280..,\\
R_1 + R_2 &\leq I(U;Y) + I(X;Z|U) = 0.3711...\\
R_1 + R_2 &\leq I(V;Z) + I(X;Y|V) = 0.3711...
\end{aligned} \tag{4.2}$$

Thus the line segment joining $(R_1, R_2) = (0.2280.., 0.1431..)$ to $(R_1, R_2) = (0.1431.., 0.2280..)$ lies on the boundary of $\mathcal{C}$. ■

The line segment joining $(R_1, R_2) = (0.2280.., 0.1431..)$ to $(R_1, R_2) = (0.1431.., 0.2280..)$ that lies on the boundary of $\mathcal{C}$ is strictly outside the line segment joining $(R_1, R_2) = (0.2411.., 0.1204..)$ to $(R_1, R_2) = (0.1204.., 0.2411..)$ that lies on the boundary of the Cover-van der Meulen region $\mathcal{D}$ [9].

Consider the following random variables (U, X).

$$\begin{aligned}
\mathrm{P}(U=0) &= 0.6372,\\
\mathrm{P}(U=1) &= 0.3628,\\
\mathrm{P}(X=1|U=0) &= 0.2465,\\
\mathrm{P}(X=1|U=1) &= 1.
\end{aligned}$$

For this pair $I(U;Y) = 0.18616..$ and $I(X;Z|U) = 0.18614...$ Hence the point $(R_1, R_2) = (0.1861, 0.1861)$ lies inside the region $\mathcal{O}_y$. By symmetry, the same point lies inside $\mathcal{O}_z$ and hence it lies inside $\mathcal{O}_y \cap \mathcal{O}_z$, the Körner-Marton outer bound. Note that $R_1 + R_2 = 0.3722 > 0.3711..$ and therefore this point lies outside $\mathcal{C}$.

5. Conclusion

We presented a new outer bound on the capacity region of the DM broadcast channel (Theorem 2.1), which is tight for all special cases where capacity is known. We then specialized the bound to the case of no common information (see (3.1)). Considering the weaker version of this bound given in Theorem 3.1, we showed that our general outer bound is strictly smaller than the Körner-Marton for the BSS channel. The outer bound in Theorem 3.1, however, is strictly larger than the Cover-van der Meulen region for this channel. We suspect that in general the outer bound in (3.1) is strictly tighter than that in Theorem 3.1. We have not been able to verify this for the BSS due to the complexity of evaluating (3.1). Finally, it would be interesting to show that our new outer bound is tight for some new class of broadcast channels that may perhaps include the BSSC.

References

[1] R F Ahlswede and J Kőrner. Source coding with side information and a converse for degraded broadcast channels. *IEEE Trans. Info. Theory*, IT-21(6):629–637, November, 1975.

[2] P F Bergmans. Coding theorem for broadcast channels with degraded components. *IEEE Trans. Info. Theory*, IT-15:197–207, March, 1973.

[3] T Cover. Broadcast channels. *IEEE Trans. Info. Theory*, IT-18:2–14, January, 1972.

[4] T Cover. An acheivable rate region for the broadcast channel. *IEEE Trans. Info. Theory*, IT-21:399–404, July, 1975.

[5] I Csizár and J Kőrner. Broadcast channels with confidential messages. *IEEE Trans. Info. Theory*, IT-24:339–348, May, 1978.

[6] A El Gamal. The capacity of a class of broadcast channels. *IEEE Trans. Info. Theory*, IT-25:166–169, March, 1979.

[7] R G Gallager. Capacity and coding for degraded broadcast channels. *Probl. Peredac. Inform.*, 10(3):3–14, 1974.
[8] S I Gelfand and M S Pinsker. Capacity of a broadcast channel with one deterministic component. *Probl. Inform. Transm.*, 16(1):17–25, Jan. - Mar., 1980.
[9] B Hajek and M Pursley. Evaluation of an achievable rate region for the broadcast channel. *IEEE Trans. Info. Theory*, IT-25:36–46, January, 1979.
[10] J Kőrner and K Marton. A source network problem involving the comparison of two channels ii. *Trans. Colloquim Inform. Theory, Keszthely, Hungary*, Auguts, 1975.
[11] K Marton. The capacity region of deterministic broadcast channels. *Trans. Int. Symp. Inform. Theory*, 1977.
[12] K Marton. A coding theorem for the discrete memoryless broadcast channel. *IEEE Trans. Info. Theory*, IT-25:306–311, May, 1979.
[13] M S Pinsker. Capacity of noiseless broadcast channels. *Probl. Pered. Inform.*, 14(2):28–334, Apr.- Jun., 1978.
[14] H Sato. An outer bound to the capacity region of broadcast channels. *IEEE Trans. Info. Theory*, IT-24:374–377, May, 1978.
[15] E van der Meulen. Random coding theorems for the discrete memoryless broadcast channel. *IEEE Trans. Info. Theory*, IT-21:180–190, March, 1975.

Chandra Nair Chandra Nair is a Post-Doctoral researcher with the theory group at Microsoft Research, Redmond. He obtained his PhD from the Electrical Engineering Department at Stanford University in June 2005. He obtained the Bachelor's degree in Electrical Engineering from IIT, Madras. His research interests are in discrete optimization problems arising in Electrical Engineering and Computer Science, algorithm design, networking and information theory. He has received the Stanford and Microsoft Graduate Fellowships (2000-2004, 2005) for his graduate studies, and he was awarded the Philips and Siemens(India) Prizes in 1999 for his undergraduate academic performance.

Abbas El Gamal Abbas El Gamal (S'71-M'73-SM'83-F'00) received his B.Sc. degree in Electrical Engineering from Cairo University in 1972, the M.S. in Statistics and the PhD in Electrical Engineering from Stanford in 1977 and 1978, respectively. From 1978 to 1980 he was an Assistant Professor of Electrical Engineering at USC. He has been on the Stanford faculty since 1981, where he is currently Professor of Electrical Engineering and the Director of the Information Systems Laboratory. He was on leave from Stanford from 1984 to 1988 first as Director of LSI Logic Research Lab, then as cofounder and Chief Scientist of Actel Corporation. In 1990 he co-founded Silicon Architects, which was later acquired by Synopsys. His research has spanned several areas, including information theory, digital imaging, and integrated circuit design and design automation. He has authored or coauthored over 150 papers and 25 patents in these areas.